

ASSESSING THE FEASIBILITY OF MTURK FOR CROSS-NATIONAL CONSUMER ONLINE SAMPLING

Nina Reynolds, University of Southampton, United Kingdom
Luke Greenacre, University of Southampton, United Kingdom

ABSTRACT

Sampling respondents from online sources is now a common research practice. With the increase in internet use across the globe, it is possible to source respondents from multiple countries. This leads to questions concerning the suitability of these samples for cross-national research. Amazon's Mechanical Turk (MTurk) is a now common source of participants in research. To assess the feasibility of MTurk for cross-national consumer research this paper asks two broad questions (i) where are MTurk workers located and can location be considered equivalent to nationality?, and (ii) are MTurk workers representative of their nationalities and/or cross-nationally comparable?

The type of consumer sample that is desired when sampling across national/cultural boundaries depends on the objectives of the research. Research objectives may require within nationality 'representativeness' or cross-national 'comparability' (Reynolds et al., 2003). It is unlikely that both representativeness and comparability will be achieved through sampling alone (Craig & Douglas, 2005). If the research objective is concerned with examining attitudes or behaviors within specific countries (descriptive international research) then the sample needs to focus on achieving representativeness (Reynolds et al., 2003). In contrast, comparability is the sampling objective when the research objective is concerned with examining the differences/similarities between consumers in different countries (comparative international research) or when looking at the cross-national generalizability of a theoretical model (theoretical international research) (Reynolds et al., 2003). With both comparative and theoretical international research, the researcher wants to isolate the impact of variables not associated with the nationality of the respondents on the results. Research objectives also point to the likely sampling issues in the cross-national environment. *Coverage* issues are particularly relevant to the question of sample representativeness as cross-national researchers need to be confident that online sampling will not disproportionately excluded any members of their target population. The general lack of *sampling frames* in the online environment is also problematic for representativeness, particularly as *non-response* errors cannot be assessed in the absence of a clear sampling frame. When comparability is the objective of the research there are fewer formal procedures for assessing sample quality, though methods do exist for selecting samples.

MTurk is "a novel, open online marketplace for getting work done by others" (Buhrmester et al., 2011, p.3); defining a survey or experiment as a 'work-task' with a payment attached allows researchers to use MTurk as a participant pool (Mason and Suri 2012). MTurk has also been used to source international respondents – for example, to investigate whether the underlying reason for some gender differences persist across nationalities (Eriksson & Simpson, 2010). The unrestricted sample associated with this study reveals that the US and India supply the majority of MTurk workers. MTurk's US-based workers have been found to provide psychological data that is reliable, and with results that are comparable to those found with traditional US samples (Berinsky et al., 2012; Goodman et al., 2012; Mason and Suri 2012; Paolacci et al., 2010).

This study's results confirmed that only the US and India have a sufficiently large pool of MTurk workers to be able to easily draw a national sample from. While almost three hundred responses were from outside the US and India, no single nationality was large enough to produce a feasible national sample. It also appears that location is a reasonable proxy for nationality with the US and India. The question of whether the obtainable national samples were (a) representative and (b) comparable indicates that it is possible to obtain an MTurk sample that is broadly representative of the general US population, but not of the Indian (urban) population. In terms of comparability, MTurk workers from India and the US are not statistically different with respect to gender. However, statistically significant differences are found with other variables. To obtain comparable US and Indian samples some control to ensure comparability would be needed. Comparable samples, as such, appear achievable.

In conclusion, when the sampling objectives of representativeness and comparability are considered, it is possible to collect a sample representative of the US population, but not the Indian (urban) population. Collecting comparable samples from India and the US using MTurk is also generally feasible. Other sources of participants will be needed if achieving the research objectives requires different nationalities. MTurk workers are, overall, a feasible sampling choice for cross-national researchers on research projects that can legitimately use the US and India as country choices, and when the research objectives are comparative or theoretical rather than descriptive.

References available on request